

# HOW CAN WE MAKE GAN PERFORM BETTER IN SINGLE MEDICAL IMAGE SUPER-RESOLUTION? A LESION FOCUSED MULTI-SCALE APPROACH

Jin Zhu<sup>\*</sup>      Guang Yang<sup>†‡</sup>      Pietro Lio<sup>\*</sup>

<sup>\*</sup> The Computer Laboratory, University of Cambridge, Cambridge, CB3 0FD, UK

<sup>†</sup> Cardiovascular Research Centre, Royal Brompton Hospital, London, SW3 6NP, UK

<sup>‡</sup> National Heart and Lung Institute, Imperial College London, London, SW7 2AZ, UK

## ABSTRACT

Single image super-resolution (SISR) is of great importance as a low-level computer vision task. The fast development of Generative Adversarial Network (GAN) based deep learning architectures realises an efficient and effective SISR to boost the spatial resolution of natural images captured by digital cameras. However, the SISR for medical images is still a very challenging problem. This is due to (1) compared to natural images, in general, medical images have lower signal to noise ratios, (2) GAN based models pre-trained on natural images may synthesise unrealistic patterns in medical images which could affect the clinical interpretation and diagnosis, and (3) the vanilla GAN architecture may suffer from unstable training and collapse mode that can also affect the SISR results. In this paper, we propose a novel lesion focused SR (LFSR) method, which incorporates GAN to achieve perceptually realistic SISR results for brain tumour MRI images. More importantly, we test and make comparison using recently developed GAN variations, e.g., Wasserstein GAN (WGAN) and WGAN with Gradient Penalty (WGAN-GP), and propose a novel multi-scale GAN (MS-GAN), to achieve a more stabilised and efficient training and improved perceptual quality of the super-resolved results. Based on both quantitative evaluations and our designed mean opinion score, the proposed LFSR coupled with MS-GAN has performed better in terms of both perceptual quality and efficiency.

**Index Terms**— Generative Adversarial Network, Super-Resolution, Medical Image Analysis, Lesion Detection, Image Processing

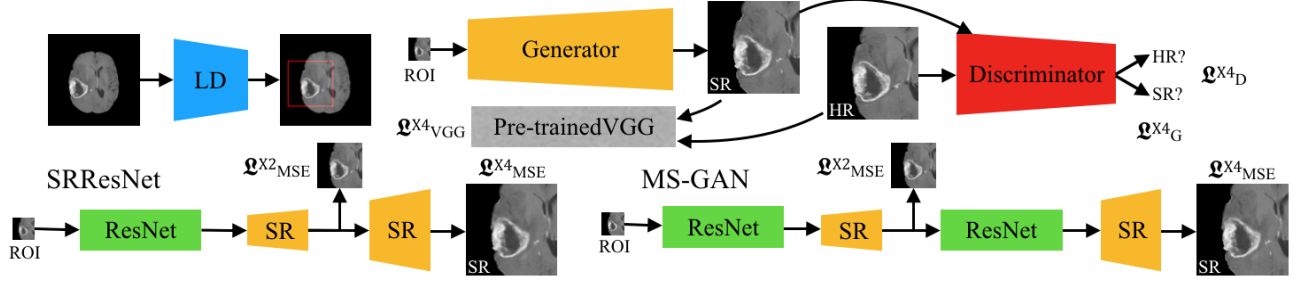
## 1. INTRODUCTION

Images with high resolution (HR) are greatly in demand for many real applications, e.g., medical images for clinical tasks, geographic information systems, security video surveillance and others [1]. However, the resolution and quality of the images are normally limited by the imaging hardware [2], effectiveness and costs. Medical images are strongly desirable with HR, because they provide crucial details of the anatomical, physiological, functional and metabolic information of patients. In addition to the potential restrictions of the imaging hardware aforementioned, medical images are more susceptible by the health limitations (e.g., ionising radiation dose of using X-ray) and acquisition time limitations (e.g., Specific Absorption Rate limits of using MRI). Moreover, movements due to patients' fatigue and organs pulsation further degrade image qualities and result in lower signal-to-noise ratio (SNR) images. Low resolution (LR) medical images with limited field of view and degraded image quality could reduce the visibility of vital pathological details and compromise the diagnostic accuracy and prognosis [3, 4].

Research studies have shown that instead of optimising hardware settings and imaging sequences, image super-resolution (SR) provides an alternative and relatively cheaper solution for spatial resolution enhancement. Compared to conventional interpolation methods, these SR methods tend to provide better SR outputs with higher SNR and less blurry effects due to the information from multiple LR images or LR-HR image pairs. Reconstruction based SR algorithms have been proven their effectiveness by recovering the HR output with fusing multiple LR images [5]. However, this type of methods is time-consuming, and the required multi-view LR images are not always available in medical image applications [6]. Learning based SR methods attract more and more attention now due to their better performance. Simply speaking, this type of SR methods learn a mapping function between LR-HR training pairs (whole images and patches), and apply this mapping to a single testing image to achieve the SR results, namely single image super-resolution (SISR) [1, 7, 8].

Recently, deep learning based SISR methods [9, 10, 11, 12, 13, 14] have boosted the performance of the super-resolved HR images mainly owe to the development of the computing power and the available big data. For example, SRGAN [14], which is developed based on a Generative Adversarial Network (GAN), has demonstrated perceptually better results compared to other deep residual network [15] based SISR methods [12, 9, 10]. More recently, Deng [13] proposed a multi-channel method for SISR which enhanced the objective and perceptual qualities separately. Lai et al. [16] incorporated a Laplacian pyramid SR network to progressively super-resolve the sub-band residuals of HR images at multiple pyramid levels. Although these GAN based methods works well on natural images, they are limited for medical images. These pre-trained models using natural images may synthesise unrealistic patterns which could affect the clinical interpretation and diagnosis. Moreover, input LR medical images with lower SNR can intrinsically undermine the performance of the GAN based methods. Thus, SISR for medical images is still an open and challenging problem [17, 18].

In this study, a novel lesion focused SISR method (LFSR) is developed to generate perceptually more realistic SR results and also avoid introducing non-existing features into the lesion area. Because the vanilla GAN architecture may suffer from unstable training and collapse mode problems, newly proposed Wasserstein GAN (WGAN) and WGAN with Gradient Penalty (WGAN-GP) are also tested and compared. Based on our findings, we proposed an advanced multi-scale GAN (MS-GAN) with LFSR to achieve a more stabilised and efficient training procedure and improved perceptual quality of super-resolved results. The validation has been done on MRI images acquired for brain tumour patients using both quantitative metrics and a designed mean opinion score (MOS).



**Fig. 1:** Our proposed lesion focused multi-scale super-resolution neural networks. LD is a pre-trained lesion detection neural network. Two architectures of generators (SRResNet and MS-GAN) are tested. Adversarial loss, MSE loss and perceptual loss are also displayed.

## 2. METHODS

### 2.1. Generative Adversarial Network (GAN)

The originally proposed vanilla GAN [19] contains a generator  $G$  and a discriminator  $D$  to be trained synchronously via competing with each other. In this work,  $G$  aims to generate as realistic as possible SR images to fool  $D$  and  $D$  aims to distinguish the SR images from real HR images that can be described as:

$$\hat{G}, \hat{D} = \min_G \max_D \mathbb{E}_{I_{HR}} [\log D(I_{HR})] + \mathbb{E}_{I_{LR}} [1 - \log D(G(I_{LR}))]. \quad (1)$$

where  $I_{LR}$  and  $I_{HR}$  are full size LR and HR images, and  $\mathbb{E}$  is the expectation of the  $D$ 's positive outputs (i.e. input is HR ground truth).

However, the vanilla GAN suffers from unstable training, collapsed mode and difficulties in hyper-parameters tuning. Thus, Wasserstein GAN (WGAN) [20] has proposed to replace the non-continuous divergence in  $\mathcal{L}_D$  with the Wasserstein-1 distance:

$$\hat{G}, \hat{D} = \min_G \max_{D \in \mathcal{D}} \mathbb{E}_{I_{HR}} [D(I_{HR})] - \mathbb{E}_{I_{LR}} [D(G(I_{LR}))] \quad (2)$$

In order to enforce the constraint of 1-Lipschitz functions  $\mathcal{D}$ , weights clipping is also introduced to ensure all the weights of the  $D$  in a compact range  $[-c, c]$ . WGAN has shown the advantages to ease the training, increase generative diversity, and promote model flexibility. However, the weights clipping was also found problematic in real applications. Thus, Gradient Penalty [21] was proposed and added into the  $\mathcal{L}_D$ , i.e., WGAN-GP.

### 2.2. Single Image Super-Resolution (SISR)

The SISR aims to find a transformation to map the distribution of the LR images  $P_{LR}$  into the distribution of the HR images  $P_{HR}$ . A dataset of LR-HR pairs is given as samples of  $P_{LR}$  and  $P_{HR}$  to derive the unknown transformation in a supervised learning manner. In this work, we use deep neural networks to estimate a generated SR distribution  $P_{SR}$ , which has the same dimension as  $P_{HR}$ , and make  $P_{SR} \approx P_{HR}$  according to the chosen metrics. However, in practice, it is challenging to achieve this goal when  $P_{SR}$  and  $P_{HR}$  are in a high-dimensional space, because it has very low possibility for them to be overlapped. Thus, we propose two strategies to tackle this problem: lesion focused SISR (LFSR) and multi-scale GAN based SISR (MS-GAN) (Fig.1).

#### 2.2.1. Lesion Focused SISR (LFSR)

Firstly, We propose LFSR, which contains a lesion detection neural network LD and a deep residual super-resolution neural network SRResNet [14]. LD aims to detect a region of interest (ROI) of

the lesions or abnormalities, e.g., brain tumours in our current study, denoted as  $I_{lr}$  and  $I_{hr}$  from the full size LR and HR images  $I_{LR}$  and  $I_{HR}$  before we applying the SRResNet, i.e.,  $I_{lr,hr} = LD(I_{LR,HR})$ .

By using LD, the LR-HR image distributions  $P_{LR}$  and  $P_{HR}$  are down-scaled as  $P_{lr}$  and  $P_{hr}$  in a lower dimensional space. Since only these ROIs are interested in clinical studies, the dimension reduction retains most of the meaningful information from the original full size images. This benefits the training of generative SR models in three aspects: (1) it significantly reduces the training cost of the GAN for the SR task, with a huge reduction of parameters need to be trained; (2) it results in SR images with better perceptual qualities via replacing the estimation of transformation from  $P_{LR}$  to  $P_{SR}$  with a much simpler one from  $P_{lr}$  to  $P_{sr}$ ; and (3) the excluded regions will not enter the training process and less artefacts will be synthesised.

#### 2.2.2. Multi-Scale GAN Based SR (MS-GAN)

The original SRResNet generates SR images by solving:

$$\hat{G} = \min_G \mathcal{L}_{SR}(G(I_{lr}), I_{hr}) \quad (3)$$

where  $\mathcal{L}_{SR}$  can be any predefined loss function. In this paper, we use pixel-wise mean-square-error (MSE) and the VGG [22] based perceptual loss:

$$\mathcal{L}_{MSE}(I_{sr}, I_{hr}) = \frac{1}{W \times H} \sum_{i,j \in W,H} (I_{sr}[i,j] - I_{hr}[i,j])^2 \quad (4)$$

$$\mathcal{L}_{VGG}(I_{sr}, I_{hr}) = \mathcal{L}_{MSE}(V_l(I_{sr}), V_l(I_{hr})) \quad (5)$$

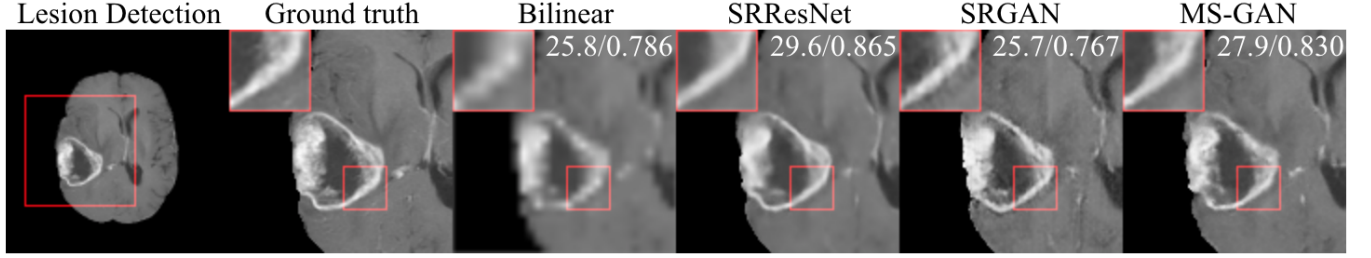
where  $W$  and  $H$  are the width and height of  $I_{sr}$  and  $I_{hr}$ , and  $V_l$  is the  $l$ th layer feature maps of the pre-trained VGG.

Since the original SRResNet generates only one scale SR images, it is hard to stabilise the optimisation process of SR tasks with higher magnifying factors (e.g. X4 magnification). Thus, we propose a MS-GAN architecture to decompose this difficult problem into a series of simpler sub-problems. Our MS-GAN (Fig. 1) can generate multi-scale SR images, and the higher dimensional images are achieved from the lower dimensional ones. For the X4 SR task, both X2 and X4 SR images are sequentially generated. Since the image quality of X4 outputs is based on the performance of X2 ones, the training procedure becomes:

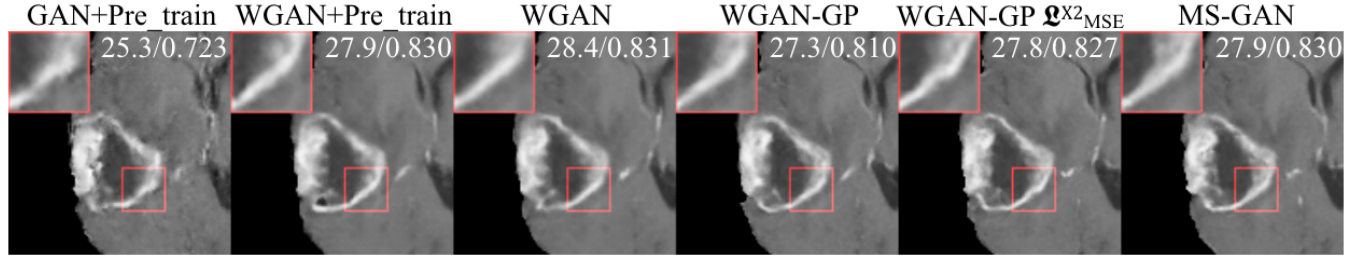
$$\hat{G} = \min_G (\mathcal{L}_{SR}^{X2}(G(I_{lr}, X2), I_{dr}) + \mathcal{L}_{SR}^{X4}(G(I_{lr}, X4), I_{hr})) \quad (6)$$

where  $I_{dr}$  is the X2 down-sampled version of the  $I_{hr}$ . In this work, we choose  $\mathcal{L}_{SR}^{X2} = \mathcal{L}_{MSE}$ , and  $\mathcal{L}_{SR}^{X4} = \mathcal{L}_{MSE} + \mathcal{L}_{VGG}$ , to avoid introducing non-realistic textures in the early stage. In addition with the adversarial loss of GAN  $\mathcal{L}_G^{X4} = -\mathbb{E}[D(G(I_{LR}))]$ , the overall loss function of our generator can be denoted as:

$$\mathcal{L}_{SR} = \mathcal{L}_{MSE}^{X2} + \mathcal{L}_{MSE}^{X4} + \mathcal{L}_{VGG}^{X4} + \mathcal{L}_G^{X4} \quad (7)$$



**Fig. 2:** The detected ROI and the predicted SR images, PSNR/SSIM are also displayed.



**Fig. 3:** The generated SR images of LFSR with GAN variations and MS-GAN. PSNR/SSIM are also displayed.

### 2.3. Data Pre-processing and Experiment Settings

The experiments have been done using the open access BraTS 2018 datasets, which contains MRI images acquired from brain tumour patients. In total, 163 patient datasets were included in our study and they were randomly divided into training (9559 slices) and independent validation (2368 slices) groups. All the images were normalised to zero-mean-unit-variance and the LR images were simulated by downsampling the HR ground truth images.

All the implementation was using Python 3.5, with TensorFlow [23] and TensorLayer, which is now widely used in solving various medical image analysis problems [24, 25, 26]. All the experiments were performed on a Linux workstation with one NVIDIA TITAN X Pascal GPU and Intel(R) Xeon(R) CPU E5-2630 v4 2.20GHz CPUs. All neural networks were trained and tested on the GPU, and the CPUs were only used for data loading and saving.

For a comparison study, we implemented and tested 6 GAN based variations. Firstly, the SRResNet [14] based LFSR with the vanilla GAN [19] was tested with a pre-training of the SRResNet to stabilise the following training of GAN (i.e., 1. GAN+Pre.train). Then, we tested the same LFSR coupled with WGAN [20], with and without the pre-training of the SRResNet (i.e., 2. WGAN+Pre.train and 3. WGAN). Furthermore, the same LFSR with WGAN-GP [21] was trained with and without  $\mathcal{L}_{MSE}^{X^2}$  as an extra term of loss function (4. WGAN-GP and 5. WGAN-GP  $\mathcal{L}_{MSE}^{X^2}$ ). Finally, we tested our proposed LFSR coupled with MS-GAN method (6. MS-GAN). All experiments used the same initial learning rate of  $10^{-4}$ , which decayed to  $10^{-5}$  at the midpoint of the training. Although WGAN and WGAN-GP based methods might converge faster than others, all tested methods were trained for 300 epochs to establish a fair comparison. In addition, we also tested the bilinear interpolation, SRResNet and SRGAN[14] for a comprehensive study.

### 2.4. Evaluation Metrics

Conventional Peak SNR (PSNR) and Structural SIMilarity (SSIM) index were used to measure the pixel-wise various and image-wise similarity between generated SR results and ground truth HR images. We also designed and performed a mean opinion score (MOS) based evaluation to quantify the perceptual reality of generated SR

images. In this study, 100 validation slices were randomly selected for MOS evaluation. For each slice, there were 1 HR ground truth and 6 SR results corresponding to the 6 GAN based variations we tested. Then, we randomly shuffled these 700 images (including 100 HR ground truths). An MR physicist (>6 years experience on brain tumour MRI images) performed blinded scoring for these shuffled images based on a Likert-type scale—0 (non-diagnostic), 1 (poor), 2 (fair), 3 (good), and 4 (great)—depending on the image qualities [27, 28]: over-smooth (S); motion and other kind of artefacts (A); unrealistic textures (U); and too noisy or low SNR (N). The MOS was then derived by calculating the mean and standard deviation for each method.

## 3. RESULTS AND DISCUSSIONS

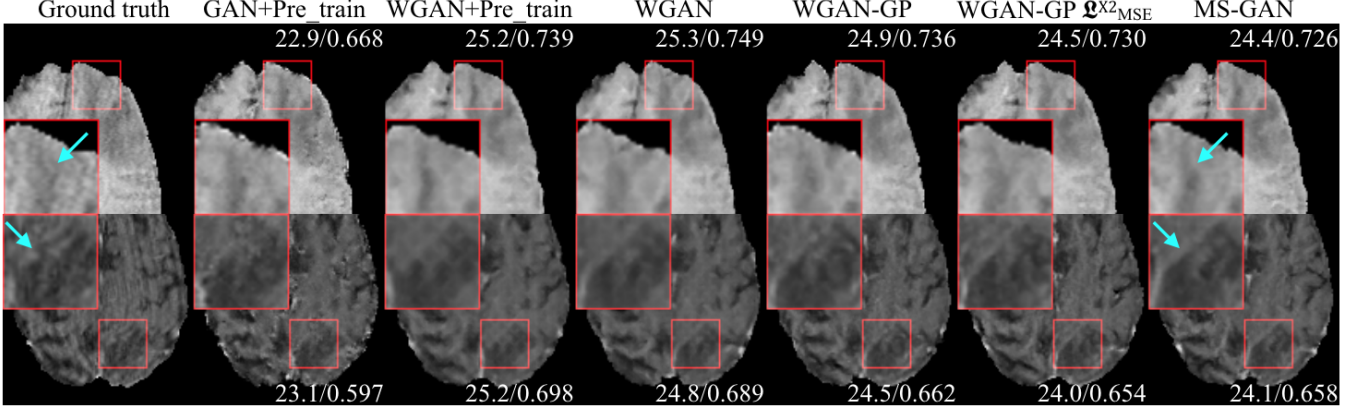
We have tested our proposed LFSR with 6 GAN variations (including the proposed MS-GAN) and different SR image generators for the X4 SR task. It is of note that in order to demonstrate the effectiveness of our MS-GAN, we showed the results of a more challenging X4 SR task, but our proposed methods can also work well with lower magnifying factors (results not shown). Table 1 tabulates the quantitative results of using LFSR coupled with different GAN models. Except the vanilla GAN produced relatively poor PSNR/SSIM, other GAN variations resulted in similar high PSNR/SSIM. Our MS-GAN method obtained the highest MOS. Figs.2 and 3 show the qualitative visualisation of an example slice. Our MS-GAN achieved high PSNR/SSIM with lesion edge and textural information preserved well. Clearly, compared to the ground truth, vanilla GAN produced noisier SR results. All WGAN based models achieved similar results, but slightly smoother than the results produced by our MS-GAN. Compared to our MS-GAN, although SRResNet yielded higher PSNR/SSIM, perceptually the results were more blurry. SRGAN achieved lower PSNR/SSIM mainly due to the synthesised stripy artefacts in the SR results with less SNR. It is of note that both SRResNet and SRGAN were applied on the whole slice but only the ROIs were evaluated (Fig. 2). All the learning based SR methods showed significant improvement over the bilinear interpolation.

We also evaluated the training and inference efficiency of all methods. The generators influenced on both training and inference



**Table 1:** Quantification comparison using PSNR, SSIM and MOS (the best *performance* in bold)

|                                   | MOS               | Poor | Fair | Good | Great | S  | A  | U  | N  | PSNR             | SSIM                |
|-----------------------------------|-------------------|------|------|------|-------|----|----|----|----|------------------|---------------------|
| GAN+Pre_train                     | 1.80±0.529        | 25   | 71   | 3    | 1     | 2  | 5  | 92 | 91 | 25.2±1.88        | 0.715±0.0670        |
| WGAN+Pre_train                    | 3.15±0.669        | 1    | 13   | 56   | 30    | 19 | 0  | 4  | 3  | 27.0±1.94        | 0.804±0.0498        |
| WGAN                              | 3.09±0.694        | 1    | 17   | 54   | 28    | 18 | 3  | 7  | 3  | <b>27.2±1.96</b> | <b>0.806±0.0499</b> |
| WGAN-GP                           | 3.10±0.686        | 2    | 13   | 58   | 27    | 3  | 0  | 14 | 14 | 26.7±1.96        | 0.788±0.0558        |
| WGAN-GP+ $\mathcal{L}_{MSE}^{X2}$ | 3.15±0.698        | 1    | 15   | 52   | 32    | 3  | 0  | 17 | 17 | 26.2±2.00        | 0.786±0.0577        |
| MS-GAN                            | <b>3.26±0.626</b> | 1    | 7    | 57   | 35    | 5  | 1  | 11 | 8  | 26.7±1.97        | 0.789±0.0554        |
| Ground Truth                      | 3.28±0.838        | 3    | 16   | 31   | 50    | 1  | 20 | 14 | 9  |                  |                     |

**Fig. 4:** GAN based methods can remove the artefacts in poor image quality ground truth images. Furthermore, our method MS-GAN can enhance the edges and textures of tumour regions. PSNR and SSIM of each result are also displayed.

costs, while the GAN variations only affected the training cost. Our LFSR with SRResNet [14] and the vanilla GAN [19] costs 229.6s/epoch for training and 4.04s to generate SR images for the whole validation dataset (2368 slices). According to additional calculation of weight clipping and gradients in WGAN [20] and WGAN-GP [21], the training time increased to 233.8s/epoch and 305.7s/epoch. Moreover,  $\mathcal{L}_{MSE}^{X2}$  also slowed down the training process slightly (314.3s/epoch using WGAN-GP). Finally, because our multi-scale SR generator has more layers, it increases both the training (422.2s/epoch) and inference costs (7.75s for the whole validation dataset). Although the SRResNet took the least cost for each training epoch, it converged much slower than all the others.

Based on our comparative study, there are several interesting findings of the GAN based models: (1) because WGAN and WGAN-GP can stabilise the training better than the vanilla GAN, the pre-training of the generator is no longer necessary; (2) both WGAN and WGAN-GP can provide perceptually more realistic SR than the vanilla GAN, and result in better PSNR/SSIM and significant improvement of the MOS; (3) our proposed LFSR coupled with MS-GAN achieved the most realistic SR with the highest MOS close to the MOS of the ground truth images.

Similar to [14], our study has also demonstrated the limitations of using PSNR/SSIM as evaluation metrics for medical image SR tasks. Although blurry images are not perceptually realistic enough, they can still achieve relatively high PSNR/SSIM. Comparing all the methods, SRResNet has achieved the highest PSNR/SSIM, but it has also smoothed out the edge and textural information of the lesion, which are useful and crucial for clinical diagnosis.

Interestingly, our proposed LFSR with MS-GAN method shows image quality improvement and signal restoration along with the SR. In Fig.4, we can observe that for these two example slices, the ground truth images are with lower SNR and obvious aliasing artefacts (thus, relatively lower MOS). Our MS-GAN method can improve the image quality by boosting the SNR and reducing the arte-

facts that has resulted in better lesion characteristics (cyan arrows in Fig. 4). We can envisage the benefits of our proposed MS-GAN based SISR method for the following clinical image analysis, segmentation and biomarker extraction and characterisation tasks.

#### 4. CONCLUSION

In this study, we propose a novel SISR method to achieve spatial resolution enhancement for the brain tumour MRI images without introducing unrealistic textures. The merits of our work are three-fold: (1) a LFSR has been developed to constrain the deep network to focus on the lesion ROIs, which does not only imitate the clinicians' scrutinization procedure, e.g., enlarge the ROIs, but also dramatically reduce the possible synthesised artefacts from the organs beyond the lesion areas; (2) a comparison study has been carried out to test vanilla GAN with newly proposed WGAN and WGAN-GP to seek possible better GAN based solutions for a more stabilised and efficient training that can yield an improved perceptual quality for the super-resolved results; (3) based on the promises of LFSR and more advanced GAN architectures, a novel MS-GAN model has been developed to tackle the challenges of SISR for medical images especially for the more tricky cases with X4 magnification. In addition to the widely used quantitative metrics (PSNR/SSIM), we also propose the MOS that incorporates experts' domain knowledge for the evaluation of the medical image SR results. Results have shown that our proposed LFSR with MS-GAN can achieve efficient SISR for brain tumour MRI images and we can envisage such models to be successfully applied for a wider range of clinical applications.

#### 5. ACKNOWLEDGEMENTS

Jin Zhu's PhD research is funded by China Scholarship Council (grant No.201708060173). Guang Yang is funded by the British Heart Foundation Project Grant (Project Number: PG/16/78/32402).

## 6. REFERENCES

- [1] Dinh Hoan Trinh, et al., “Novel example-based method for super-resolution and denoising of medical images,” *IEEE Trans. Image Processing*, vol. 23, no. 4, pp. 1882–1895, 2014.
- [2] William W Moses, “Fundamental limits of spatial resolution in pet,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 648, pp. S236–S240, 2011.
- [3] Guang Yang, et al., “Combined self-learning based single-image super-resolution and dual-tree complex wavelet transform denoising for medical images,” in *Medical Imaging 2016: Image Processing*. International Society for Optics and Photonics, 2016, vol. 9784, p. 97840L.
- [4] Guang Yang, et al., “Super-resolved enhancement of a single image and its application in cardiac mri,” in *International Conference on Image and Signal Processing*. Springer, 2016, pp. 179–190.
- [5] Chih-Yuan Yang, Jia-Bin Huang, and Ming-Hsuan Yang, “Exploiting self-similarities for single frame super-resolution,” in *Asian conference on computer vision*. Springer, 2010, pp. 497–510.
- [6] Li-Wei Kang, et al., “Self-learning-based single image super-resolution of a highly compressed image,” in *Multimedia Signal Processing (MMSP), 2013 IEEE 15th International Workshop on*. IEEE, 2013, pp. 224–229.
- [7] Hong Chang, Dit-Yan Yeung, and Yimin Xiong, “Super-resolution through neighbor embedding,” in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*. IEEE, 2004, vol. 1, pp. I–I.
- [8] Jianchao Yang, et al., “Coupled dictionary training for image super-resolution,” *IEEE transactions on image processing*, vol. 21, no. 8, pp. 3467–3478, 2012.
- [9] Ying Tai, Jian Yang, and Xiaoming Liu, “Image super-resolution via deep recursive residual network,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, vol. 1, p. 5.
- [10] Bee Lim, et al., “Enhanced deep residual networks for single image super-resolution,” in *The IEEE conference on computer vision and pattern recognition (CVPR) workshops*, 2017, vol. 1, p. 4.
- [11] Zheng Hui, Xiumei Wang, and Xinbo Gao, “Fast and accurate single image super-resolution via information distillation network,” in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.
- [12] Yulun Zhang, et al., “Residual dense network for image super-resolution,” in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.
- [13] Xin Deng, “Enhancing image quality via style transfer for single image super-resolution,” *IEEE Signal Processing Letters*, vol. 25, no. 4, pp. 571–575, 2018.
- [14] Christian Ledig, et al., “Photo-realistic single image super-resolution using a generative adversarial network,” in *CVPR*, 2017, vol. 2, p. 4.
- [15] Kaiming He, et al., “Deep residual learning for image recognition,” *arXiv preprint arXiv:1512.03385*, 2015.
- [16] Wei-Sheng Lai, et al., “Fast and accurate image super-resolution with deep laplacian pyramid networks,” *IEEE transactions on pattern analysis and machine intelligence*, 2018.
- [17] Liang Han and Zhaozheng Yin, “A cascaded refinement gan for phase contrast microscopy image super resolution,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2018, pp. 347–355.
- [18] Yuhua Chen, et al., “Efficient and accurate mri super-resolution using a generative adversarial network and 3d multi-level densely connected network,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2018, pp. 91–99.
- [19] Ian Goodfellow, et al., “Generative adversarial nets,” in *Advances in neural information processing systems*, 2014, pp. 2672–2680.
- [20] Martin Arjovsky, Soumith Chintala, and Léon Bottou, “Wasserstein gan,” *arXiv preprint arXiv:1701.07875*, 2017.
- [21] Ishaan Gulrajani, et al., “Improved training of wasserstein gans,” in *Advances in Neural Information Processing Systems*, 2017, pp. 5767–5777.
- [22] Karen Simonyan and Andrew Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.
- [23] Martín Abadi, et al., “TensorFlow: Large-scale machine learning on heterogeneous systems,” 2015, Software available from tensorflow.org.
- [24] Hao Dong, et al., “Automatic brain tumor detection and segmentation using u-net based fully convolutional networks,” in *Annual Conference on Medical Image Understanding and Analysis*. Springer, 2017, pp. 506–517.
- [25] Simiao Yu, et al., “Deep de-aliasing for fast compressive sensing mri,” *arXiv preprint arXiv:1705.07137*, 2017.
- [26] Guang Yang, et al., “Dagan: Deep de-aliasing generative adversarial networks for fast compressed sensing mri reconstruction,” *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1310–1321, 2018.
- [27] Guang Yang, et al., “Fully automatic segmentation and objective assessment of atrial scars for long-standing persistent atrial fibrillation patients using late gadolinium-enhanced mri,” *Medical physics*, vol. 45, no. 4, pp. 1562–1576, 2018.
- [28] Maximilian Seitzer, et al., “Adversarial and perceptual refinement for compressed sensing mri reconstruction,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2018, pp. 232–240.